Polycomb eviction as a new distant enhancer function

RESEARCH COMMUNICATION

Polycomb eviction as a new distant enhancer function

Douglas Vernimmen,[1] Magnus D. Lynch,[1] Marco De Gobbi,[1] David Garrick,[1] Jacqueline A. Sharpe,[1] Jacqueline A. Sloane-Stanley,[1] Andrew J.H. Smith,[1,2] and Douglas R. Higgs[1,3]

[1]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, United Kingdom; [2]Institute for Stem Cell Research, University of Edinburgh, Edinburgh EH9 3JQ, United Kingdom

Remote distal enhancers may be located tens or thousands of kilobases away from their promoters. How they control gene expression is still poorly understood. Here, we analyze the influence of a remote enhancer on the balance between repression (Polycomb—PcG) and activation (Trithorax—TrxG) of a developmentally regulated gene associated with a CpG island. We reveal its essential, nonredundant role in clearing the PcG complex and H3K27me3 from the CpG island. In the absence of the enhancer, the H3K27me3 demethylase (JMJD3) is not recruited to the CpG island. We propose a new role of long-range regulatory elements in removing repressive PcG complexes.

Supplemental material is available for this article.

Received May 9, 2011; revised version accepted July 5, 2011.

It is still not fully understood how genes are switched on or off during differentiation and development. However, the role of remote regulatory elements in this process is thought to be critical (Bulger and Groudine 2011). One important epigenetic mechanism regulating many genes involves the opposing effects of the Polycomb (PcG, repressive) and Trithorax group (TrxG, activating) proteins, which play critical roles in stem cell biology, development, and cancer (Simon and Kingston 2009). The promoters of most PcG/TrxG target genes are associated with CpG islands that remain unmethylated in all cell types (Mendenhall et al. 2010; MD Lynch, AJ Smith, M De Gobbi, M Flenley, JR Hughes, D Vernimmen, H Ayyub, JA Sharpe, JA Sloane-Stanley, L Sutherland, et al., in prep.). During lineage commitment and differentiation, PcG complexes are cleared from silenced genes and replaced by TrxG complexes as these genes are switched on. Specific histone methyltransferases in these complexes create characteristic chromatin signatures when PcG (H3K27me3) or TrxG (H3K4me3) complexes are bound to chromatin. Bivalent chromatin patterns (H3K27me3 and H3K4me3) may be seen prior to commitment, and these two chromatin marks are then resolved during differentiation (Sawarkar and Paro 2010; Surface et al. 2010). Bivalent domains may also arise de novo during differentiation (Roh et al. 2006; Mohn et al. 2008). The recruitment and removal of PcG and TrxG complexes are a dynamic process. However, the potential of remote *cis*-acting elements to influence the balance between repression and activation throughout development and differentiation has not been explored. To address these issues, we characterized changes in the recruitment of transcription factors (TFs), cofactors, chromatin-associated proteins (including PcG and TrxG), and associated chromatin modifications to the human α-globin CpG island in the presence or absence of a distant enhancer.

A major problem in analyzing these aspects of gene regulation arises when perturbation of the gene in question (e.g., a key TF) leads to a perturbation of cell fate. In this case, the cellular phenotype alters as regulation of the gene under investigation is altered. To circumvent this, such experiments are often done in cell lines whose fate is fixed. To analyze human α-globin expression in primary cells, we developed an experimental model in which the human α-globin cluster (117 kb) replaces the mouse α-globin cluster (87 kb; humanized model) (Fig. 1; Wallace et al. 2007). Since α-globin is an end product of a gene regulatory network, perturbing this gene has no effect on cell fate. Expression of the human α-globin genes is controlled by remote upstream regulatory elements (MCSR1–4), of which MCS-R2 (also known as HS-40) (Fig. 1), located 60 kb upstream of the α-globin promoters, is the major enhancer element (Higgs et al. 2008). In the humanized system in which MCS-R2 has been deleted by homologous recombination, the erythroid progenitors in adult mice differentiate normally, and thus, in terms of the *trans*-acting environment, the normal and mutant cells are directly comparable (Vernimmen et al. 2009). Any differences are solely due to the presence or absence of MCS-R2 in *cis* to the promoter.

Here we show that, in erythroid cells, the remote tissue-specific enhancer (MCS-R2) plays an essential, nonredundant role in clearing the PcG complex and its associated modification (H3K27me3) from the CpG island, when the α-globin genes become fully activated. Furthermore, in the absence of the enhancer, the CpG island does not recruit the H3K27 demethylase JMJD3. This demonstrates that in addition to the recruitment of TFs, cofactors, and the preinitiation complex (PIC) at the promoter, long-range regulatory elements also play a critical role in removing repressive PcG complexes.

Results and Discussion

Using the humanized mouse, we previously showed that the removal of MCS-R2 dramatically reduces the formation of chromosomal looping, which normally occurs between the upstream regulatory elements and the promoter in erythroid cells. In the absence of MCS-R2, the human α-globin genes are expressed at <2% of normal, in erythroid cells (Vernimmen et al. 2009). Here, we first determined the effect of MCS-R2 on recruitment of key, cell type-specific, and ubiquitously expressed TFs known

[*Keywords*: polycomb; enhancer; chromatin; CpG islands]
[3]Corresponding author.
E-mail doug.higgs@imm.ox.ac.uk.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.16985411. Freely available online through the *Genes & Development* Open Access option.

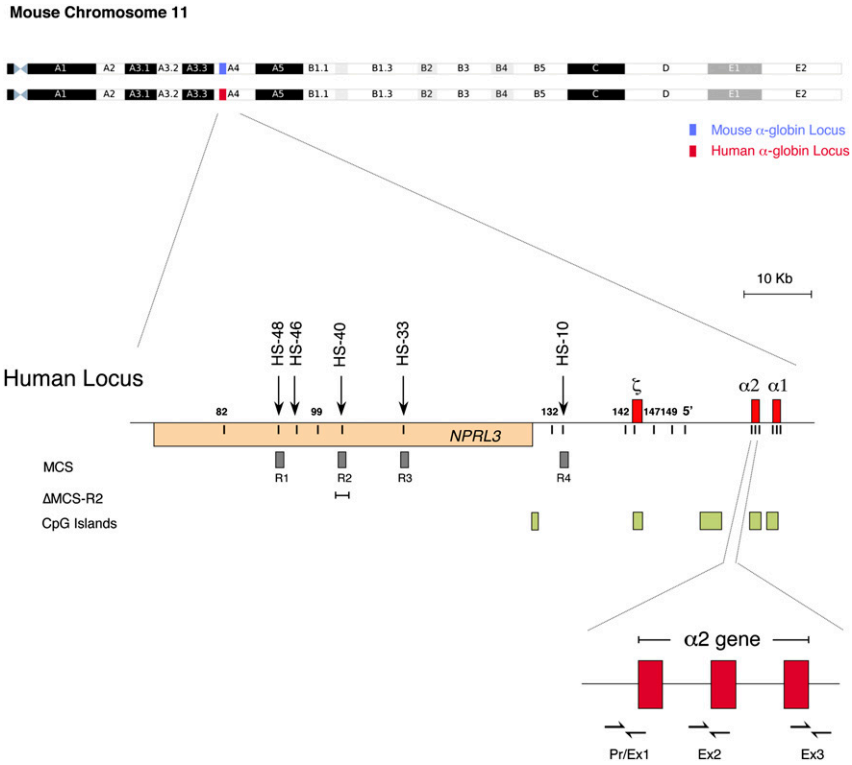

Figure 1. Generation of humanized mice as a model to study the human α-globin locus and organization of the human α-globin cluster. A region of conserved synteny in mouse chromosome 11 was replaced by the orthologous human region (Wallace et al. 2007). In this study, heterozygous mice were used to avoid lethality using the MCS-R2 deletion (ΔMCS-R2) but also to keep the other allele (mouse) as an endogenous positive control. The globin genes are shown as labeled red boxes. The positions of previously described, DNaseI-hypersensitive sites are shown as arrows. The widely expressed gene *NPRL3* (also known as the −14 gene in human or the Prox gene in mouse) transcribed from the opposite strand to that of the α-globin is shown as an orange box. Short vertical lines (black) indicate amplicons used in ChIP experiments. Gray boxes refer to previously defined multispecies conserved elements (MCS) and light-green boxes indicate CpG islands.

to bind the α-globin promoter in Ter119$^+$ mature erythroblasts (Vernimmen et al. 2007). In the absence of MCS-R2, TFs (e.g., Sp-XKLF and NFY) and cofactors, such as histone acetyl transferases (PCAF, GCN5, CBP, and p300), are not recruited efficiently (Supplemental Table 1). These findings are reflected by suboptimal histone acetylation (Supplemental Table 1). Of interest, there is a progressive decrease in TF occupancy at key *cis* elements (MCS-R1/HS-48, MCS-R3/HS-33, and MCS-R4/HS-10) along the chromosome, with the α-globin promoter being the most affected (Supplemental Fig. 1; Supplemental Table 1). Furthermore, in the absence of MCS-R2, ubiquitination of H2B (H2Bub, associated with activation) does not occur efficiently at the α-globin gene (Fig. 2, left panel). Ubiquitination of H2B is thought to be required for subsequent methylation of H3K79 by Dot1 (Shilatifard 2008), and without MCS-R2, H3K79 is not efficiently methylated (Fig. 2; Supplemental Fig. 2). As previously described (Vernimmen et al. 2007), without MCS-R2, occupancies of the PIC components (TBP, TFIIE, and TFIIB) were all low with an even more dramatic effect on PolII (Fig. 2; Supplemental Fig. 1) compared with the normal chromosome. These findings confirm that in the humanized mouse cells (as in primary human and mouse cells), a key role of the upstream element (MCS-R2) is to facilitate recruitment and/or stabilization of the PIC at the α-globin promoter, late during erythropoiesis.

We then evaluated the role of MCS-R2 on the deposition of methyl groups to the Lys 4 residue of histone H3. As in human primary erythroblasts (De Gobbi et al. 2007), the upstream elements are marked by mono- and dimethyl H3K4 in fully active humanized erythroid cells (Supplemental Fig. 3). In contrast, the chromatin associated with the α-globin promoter and the body of the gene is modified by H3K4me3 (Fig. 2, right panel; Supplemental Fig. 3). These findings are also consistent with the chromatin features of enhancers and promoters, as established in genome-wide analyses (Heintzman et al. 2007). In the absence of MCS-R2, the mono- and dimethylation of H3K4 seem to occur normally (Supplemental Fig. 3, left panel). Moreover and more unexpectedly, despite the absence of detectable H2Bub and H3K79me3 and the associated transcription (expression of only ~2% of normal), the overall level of H3K4me3 at the α-globin gene was ~50% of that seen in the fully active, highly transcribed gene (considering the amount of H3K4me3 detected on the endogenous control) (Fig. 2; Supplemental Fig. 3). Note that, in the absence of MCS-R2, the distribution of mono-, di-, and trimethyl H3K4 becomes skewed more to the promoter than the body of the gene. This might be partially due to higher levels of H3 occupancy at the promoter in the ΔMCS-R2 compared with normal (Supplemental Fig. 4). Of interest, in the absence of the enhancer, H3K36me3 (a marker of transcriptional elongation) is also detectable (Supplemental Fig. 5).

The H3K4me3 modification is imposed by the human TrxG proteins, which include several complexes (hSet1 and MLL1-5) (Shilatifard 2008). It has recently been shown that H3K4 methylation at CpG islands may be mediated via the hSet1 complex, which contains a CpG island-binding protein (CGBP, or Cfp1). CGBP is thought to bind unmethylated CpG-rich sequences irrespective of transcription, thereby recruiting hSet1 (Thomson et al. 2010). This implies that chromatin at unmethylated CpG islands might be modified by H3K4me3 as a default state. Here we found that in human nonerythroid cells, CGBP/Cfp1 is not detectable at the associated CpG island (as judged by two different antibodies) (Supplemental Fig. 6). Similarly, when the α genes are severely down-regulated in the humanized mouse MCS-R2 mutant, CGBP/Cfp1 is present at very low levels (Fig. 2; Supplemental Fig. 7). Thus, CGBP/Cfp1 is only readily detected in association with high levels of α-globin expression. Therefore, it seems unlikely that the relatively high levels of H3K4me3 seen in the ΔMCS-R2 mutant (when the α genes are expressed at relatively low levels) reflect a default state mediated by CGBP/Cfp1. In contrast, it suggests

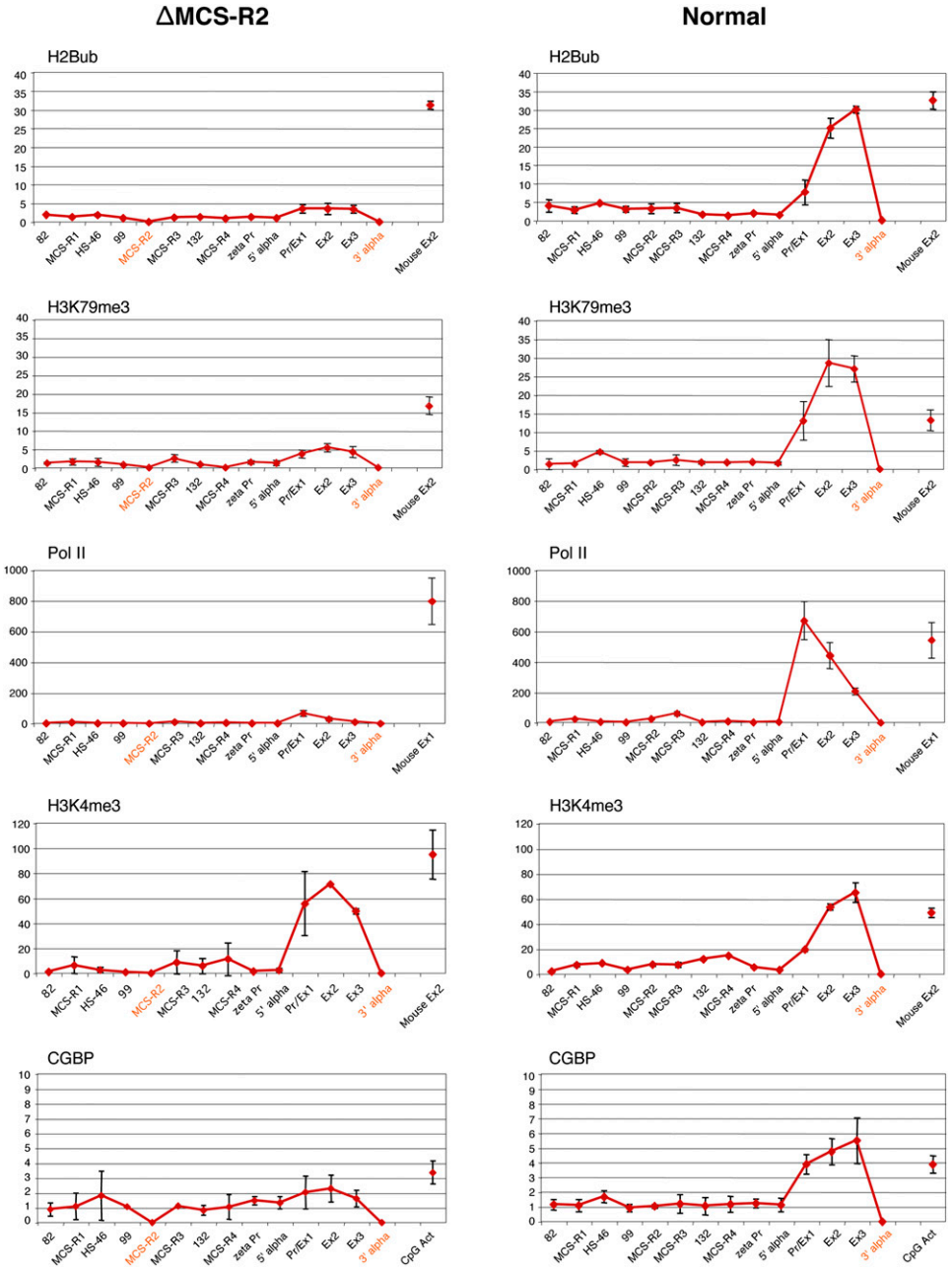

Figure 2. H3K4me3 is not associated with H2Bub, H3K79me3, Pol II, and CGBP. Real-time PCR analysis of immunoprecipitated chromatin using the antibodies indicated in −MCS-R2 (*left*) and +MCS-R2 (*right*) humanized erythroid cells (Ter119-positive cells purified by automagnetic-activated cell sorting). The *Y*-axis represents enrichment over the input DNA, normalized to a control sequence in the mouse GAPDH gene. The *X*-axis represents the positions of the TaqMan probes used. The coding sequence is represented by the three exons (Promoter/Ex1, Ex2, Ex3) of the human α-globin genes as shown in Figure 1. (Mouse Ex1 and Mouse Ex2) Control sequences at the mouse α-globin gene; (CpG Act) additional control sequence at the CpG island of the mouse β-actin gene. The amplicons highlighted in red represent deleted regions in the humanized mice, for which no PCR signal is observed. Error bars correspond to ±1 SD from at least two independent ChIPs. The CGBP antibody used here was purchased from Abcam.

that H3K4me3 might be a sensitive mark of basal transcription. Interestingly, it was previously suggested that CpG islands marked by H3K27me3 may exclude binding of CGBP (Thomson et al. 2010). We showed previously that a functionally repressive PcG complex is bound at the CpG islands associated with the human α-globin promoters in human nonerythroid cells, in which the α genes are transcriptionally silent (Garrick et al. 2008). Here, we show that CGBP is indeed excluded from CpG islands bound by PcG by comparing human nonexpressing (PcG-bound) versus expressing (PcG-unbound) cells (Supplemental Fig. 6). We therefore next determined whether the lack of CGBP recruitment in the ΔMCS-R2 mutant is associated with persistent PcG binding.

PcG complex is normally bound to both the adult (α) and embryonic (ζ) globin CpG islands in human embryonic

stem cells and pluripotent progenitors, in which expression of the α genes is detectable at extremely low levels (De Gobbi et al. 2011). Polycomb remains bound to the silenced embryonic (ζ) gene in mature erythroid cells. In contrast, PcG is cleared from the adjacent α-globin CpG island as this gene becomes expressed at high levels in mature erythroid cells (Garrick et al. 2008; De Gobbi et al. 2011). This pattern is fully recapitulated in the humanized mouse system (Fig. 3, right panel; Supplemental Fig. 8). Moreover, we show here that as the PRC2-associated chromatin modification (H3K27me3) is erased, the histone H3K27 demethylase JMJD3 is recruited (Fig. 3, right panel; Supplemental Fig. 8).

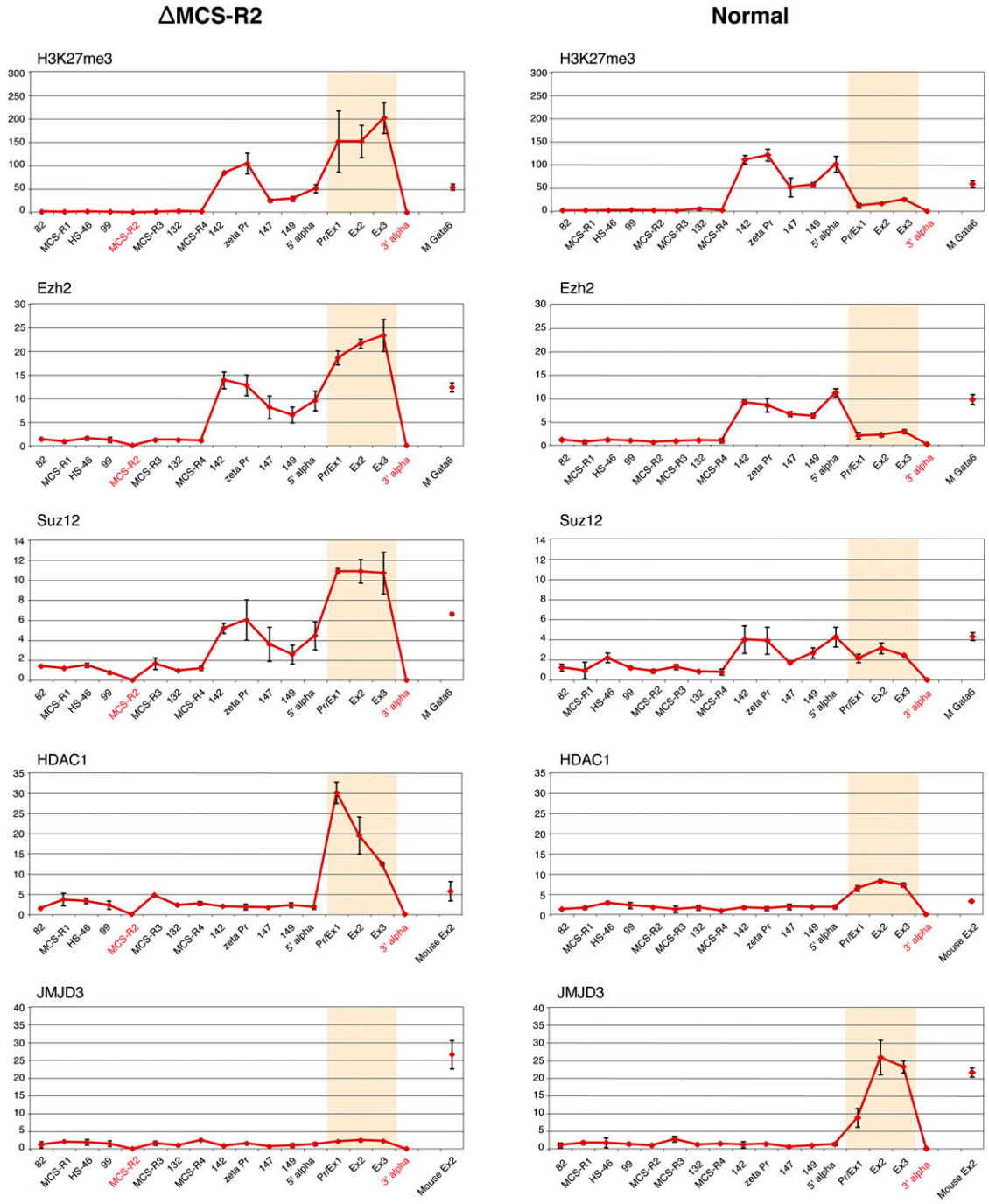

Figure 3. Removal of Polycomb repressor complexes is dependent on MCS-R2 and associated with recruitment of JMJD3. Graphs are displayed as in Figure 2. Additional amplicons flanking the embryonic ζ-globin gene (142, 147, and 149) have been used in these experiments to highlight the two independent PcG domains observed in the mutant human allele ΔMCS-R2 (shown in the *left* panel). The shaded area represents the adult α-globin CpG island, from which PcG is normally cleared in the Normal allele in erythroid cells (shown in the *right* panel). (*Left*) In the absence of MCS-R2, PcG is not cleared from the α-globin genes and thus remains present at both domains. M Gata6 represents a positive control for PcG binding. The JMJD3 antibody used recognizes the N-terminal region of the protein (Millipore).

We next studied humanized mice in which MCS-R2 had been deleted to determine whether the upstream enhancer normally plays a role in clearing the PcG complex from the α-globin CpG island during erythropoiesis. As before, in the absence of MCS-R2, the silenced ζ gene remains bound by PcG throughout differentiation. However, in the absence of MCS-R2, PRC2 is no longer cleared from the α-globin CpG island, and JMJD3 is no longer recruited (Fig. 3, left panel; Supplemental Fig. 8). We showed previously that PcG repression of the α-globin gene may be mediated (at least in part) by HDAC1, which is normally cleared with PcG complex during erythropoiesis (Garrick et al. 2008). Here, in the absence of MCS-R2, HDAC1 remains bound to the α-globin genes (Fig. 3). These contrasting observations on two adjacent CpG islands show that MCS-R2 exerts a specific effect on PcG clearance from the α-globin (but not the ζ-globin) gene during adult erythropoiesis.

There is increasing evidence that CpG islands, such as those associated with the α-globin promoter, constitute at least one element that can mediate recruitment of PcG and TrxG complexes to mammalian promoters (Mendenhall et al. 2010; MD Lynch, AJ Smith, M De Gobbi, M Flenley, JR Hughes, D Vernimmen, H Ayyub, JA Sharpe, JA Sloane-Stanley, L Sutherland, et al., in prep.). As previously noted, PcG-binding sites are dynamic, are nucleosome-depleted, and have a rapid histone turnover (the residency time of PcG is in the order of a few minutes) (Ficz et al. 2005). PcG binding is therefore thought to be dynamic and sensitive to the antagonistic action of TrxG proteins together with positive and negative input from other TFs and cofactors. However, it is not known whether the eviction of PcG silencing complex from its targets, seen during development and differentiation, depends on the presence of distal regulatory elements or only on (co)factors acting at proximal *cis* elements. In this study, we used a mouse experimental model to analyze the CpG island associated with the human α-globin promoter in two states: without and with its interacting distant enhancer, both in terminally differentiated erythroid cells. We also compared the recruitment of CGBP in nonexpressing versus expressing human cells. In nonerythroid cells, the unmethylated, nuclease-insensitive CpG island associated with the α-globin gene is bound by PcG and is transcriptionally silenced (referred to as the "silent state") (Fig. 4A). In erythroid cells without MCS-R2 (referred to as "basal state"), and in contrast to nonerythroid cells (Garrick et al. 2008), the promoter becomes accessible to some TFs and is associated with some active chromatin modifications (e.g., H3K4me3) with relatively low levels of transcription (~2% of normal) (Fig. 4B). Nevertheless, the PcG complex with its associated modification (H3K27me3) is still prominent at the α-globin CpG island. We also demonstrate here that PcG and CGBP binding are mutually exclusive (cf. Figs. 2 and 3). In erythroid cells with MCS-R2 (referred to as "active state"), PcG complexes are completely removed from the CpG island (Fig. 4C). Furthermore, the histone H3K27 demethylase JMJD3, which may remove H3K27me3 and thereby facilitate transcription, is also recruited at high levels. Recruitment of the SAGA complex (e.g., PCAF and GCN5) becomes prominent and the downstream effects (e.g., deposition of H2Bub and H3K79 methylation) are established. At this stage, high levels of transcription are associated with binding of CGBP. We thus show that the recruitment of the demethylase JMJD3 and full clearance of the PcG-repressive complex (including PRC2 and HDAC1) at the α-globin CpG island depend on one or more activities mediated by the remote regulatory element and are associated

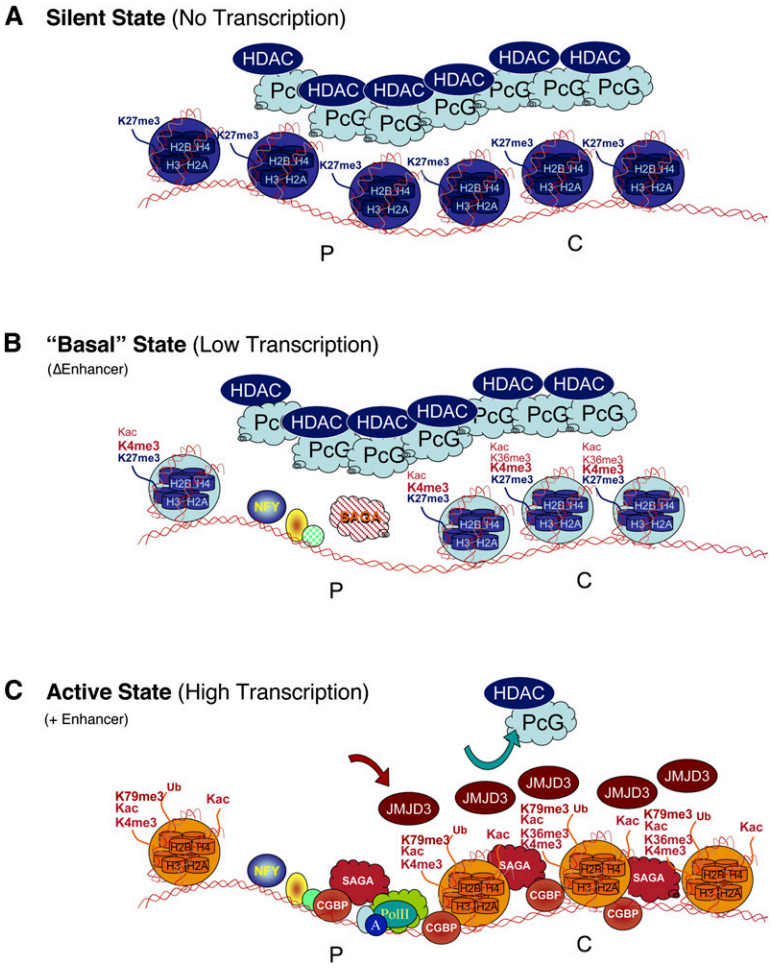

Figure 4. Long-range control of epigenetic regulation. (*A*) In nonerythroid cells, the CpG island is entirely silenced by PcG and HDAC1, associated with the repressive histone mark H3K27me3. The promoter "P" is not sensitive to DNaseI, and transcription does not occur. (*B*) In erythroid cells lacking the enhancer, the gene remains repressed by PcG and marked by H3K27me3. At this basal level of expression, the promoter becomes accessible to some TFs and chromatin-modifying enzymes and is marked by moderate levels of H3K4me3, which reflect very low levels of transcription. (*C*) In the presence of the enhancer, PcG is evicted and the H3K27me3 histone mark is erased by recruitment of demethylase JMJD3. Acetylation (H3ac and H4ac), H3K79me3, and H2Bub increases with spreading of HAT and Bre (SAGA) along the coding sequence "C." At this activated stage, the remaining TFs, including Pol II, are now fully recruited, and a high rate of transcription occurs. The CpG island at this stage is also bound by CGBP.

with the transition between basal and fully activated transcription.

These findings demonstrate for the first time that the pattern of PcG binding at a CpG island may be affected by *cis*-acting elements located far away from the associated promoter. In contrast, the chromatin modification associated with TrxG activity (H3K4me3) appears to be more dependent on local changes at the CpG island that occur in the context of basal transcription. Future studies will address how long-range enhancers exert these effects. It is possible that transcriptional activation per se competes with the competitive binding of PcG complexes and is responsible for the clearance of these complexes (MD Lynch, AJ Smith, M De Gobbi, M Flenley, JR Hughes, D Vernimmen, H Ayyub, JA Sharpe, JA Sloane-Stanley, L Sutherland, et al., in prep.). The second is that upstream elements also deliver new proteins (e.g., JMJD3) or modify proteins (e.g., histones) that facilitate the removal of PcG. In the past, detailed analysis of the globin genes has established many of the general principles underlying mammalian gene regulation, and it therefore seems probable that this new role of distal regulatory elements in removing PcG from their target promoters will be of considerable general importance.

Materials and methods

Primary cells

Ter119-positive mature primary mouse erythroid cells (humanized) were obtained by automagnetic-activated cell sorting, as previously described (Vernimmen et al. 2007, 2009; Wallace et al. 2007). Primary human erythroblasts were obtained from peripheral blood mononuclear cells (PBMCs) collected from blood donors and expanded in a two-phase system as previously described (De Gobbi et al. 2007). Epstein-Barr virus (EBV)-transformed lymphoblastoid cell lines were derived from healthy subjects.

Chromatin immunoprecipitation assay (ChIP)

ChIP was performed as previously described (Vernimmen et al. 2007). For histone-modifying enzymes, chromatin was first cross-linked with EGS (Pearce, Thermo Scientific, product no. 21565) in PBS at a final concentration of 2 mM for 60 min at room temperature. Formaldehyde was then added at a final concentration of 1% for 15 min at room temperature, and samples were sonicated for 20 min at 4°C to fragment genomic DNA (Bioruptor; Diagenode). The antibodies used were Pol II (H-224), JMJD3 (N-24), GCN5 (H-75), PCAF (H-369), p300 (C-20), CBP (A-22) (purchased from Santa Cruz Biotechnology); H3K79me1 (ab2886), H3K79me2 (ab3594), H3K79me3 (ab2621), H3K27ac (ab4729), H3 (ab1791), CGBP (ab56035), H3K4me3 (ab8085), and Suz12 (ab12073) (purchased from abcam); DOT1L (A300-954A) (purchased from Bethyl Laboratory, Inc.); HDAC1 (06-720), H3ac (06-599), H4ac (06-866), H3K9ac (07-352), H3K14ac (07-353), H3K27me3 (07-449), H3K4me1 (07-436), H3K4me2 (07-030), H3K4me3 (05-745R), H4 (07-108), H2A (06-13923), H2B (06-1790), JMJD3 (07-1533), and JMJD3 (07-1534) (purchased from Millipore); H2Bub (MM-0029) (purchased from Medimabs); and Ezh2 (PAB0649) (purchased from Abnova). Real-time PCR using primers and probes (5′FAM–3′TAMRA) for murine and human α-globin locus were described previously (Anguita et al. 2004; De Gobbi et al. 2007).

Acknowledgments

We are very grateful to Dave Skalnik for CGBP antibody. Many thanks to Jane Mellor, Tom Milne, Nick Proudfoot, Roger Patient, Maggie Walmsley, and Richard Gibbons for critically reading the manuscript. Furthermore, we thank Sue Butler and Christina Rode for technical assistance. This work was funded by the Medical Research Council and the National Institute for Health Research (NIHR) Biomedical Research Centre Program.

References

Anguita E, Hughes J, Heyworth C, Blobel GA, Wood WG, Higgs DR. 2004. Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J* **23:** 2841–2852.

Bulger M, Groudine M. 2011. Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144:** 327–339.

De Gobbi M, Anguita E, Hughes J, Sloane-Stanley JA, Sharpe JA, Koch CM, Dunham I, Gibbons RJ, Wood WG, Higgs DR. 2007. Tissue-specific histone modification and transcription factor binding in α globin gene expression. *Blood* **110:** 4503–4510.

De Gobbi M, Garrick D, Lynch M, Vernimmen D, Hughes JR, Goardon N, Sidinh L, Lower KM, Sloane-Stanley JA, Pina C, et al. 2011. Generation of bivalent chromatin domains during cell fate decisions. *Epigenetics Chromatin* **4:** 9. doi: 10.1186/1756-8935-4-9.

Ficz G, Heintzmann R, Arndt-Jovin DJ. 2005. Polycomb group protein complexes exchange rapidly in living *Drosophila*. *Development* **132:** 3963–3976.

Garrick D, De Gobbi M, Samara V, Rugless M, Holland M, Ayyub H, Lower K, Sloane-Stanley J, Gray N, Koch C, et al. 2008. The role of the polycomb complex in silencing α-globin gene expression in nonerythroid cells. *Blood* **112:** 3889–3899.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39:** 311–318.

Higgs DR, Vernimmen D, Wood B. 2008. Long-range regulation of α-globin gene expression. *Adv Genet* **61:** 143–173.

Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, Ku M, Bernstein BE. 2010. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet* **6:** e1001244. doi: 10.1371/journal.pgen.1001244.

Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D. 2008. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* **30:** 755–766.

Roh TY, Cuddapah S, Cui K, Zhao K. 2006. The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci* **103:** 15782–15787.

Sawarkar R, Paro R. 2010. Interpretation of developmental signaling at chromatin: the Polycomb perspective. *Dev Cell* **19:** 651–661.

Shilatifard A. 2008. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr Opin Cell Biol* **20:** 341–348.

Simon JA, Kingston RE. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* **10:** 697–708.

Surface LE, Thornton SR, Boyer LA. 2010. Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell* **7:** 288–298.

Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD, et al. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464:** 1082–1086.

Vernimmen D, De Gobbi M, Sloane-Stanley JA, Wood WG, Higgs DR. 2007. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J* **26:** 2041–2051.

Vernimmen D, Marques-Kranc F, Sharpe JA, Sloane-Stanley JA, Wood WG, Wallace HA, Smith AJ, Higgs DR. 2009. Chromosome looping at the human α-globin locus is mediated via the major upstream regulatory element (HS-40). *Blood* **114:** 4253–4260.

Wallace HA, Marques-Kranc F, Richardson M, Luna-Crespo F, Sharpe JA, Hughes J, Wood WG, Higgs DR, Smith AJ. 2007. Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. *Cell* **128:** 197–209.